SPHENIX DAQ/TRIGGER

Martin L. Purschke, BNL

Current PHENIX DAQ coordinator

Member of the PHENIX Online Crew since… forever (1996)

- DAQ for sPHENIX
- Trigger System
- Trigger detectors (briefly)

Some Remarks…

- PHENIX does not have (and never had) distinct “online” and “offline” groups
- ChrisP, Mike, myself… all work on “both sides of the fence”
- There is only one version of the “Event libraries” which gets you access to raw data
- For the longest time, the PHENIX offline framework was the online framework (just leave the online capabilities alone)
- Prevents any incentive to sweep stuff over the demarcation line
- And makes for a very engaged and “plugged-in” team

DAQ – the Plan

- As much as possible, we will re-use the existing PHENIX DAQ
- We have been asked by review committee members why we wouldn't want to upgrade a 15yr old DAQ
- Answer: We already did that, we have just the system that we would build today
- Legacy gear from older detectors goes out the window
- Working and very modern front-end, lots of experience
- We have traditionally taken the highest data rates in the field
- LHC-era data rates and volumes a decade before the LHC turn-on
- 15KHz event rate envisioned, 14 have been demonstrated

Some Highlights…

chep 06

LHC-Era Data Rates in 2004 and 2005 Experiences of the PHENIX Experiment with a PetaByte of Data

Martin L. Purschke, Brookhaven National Laboratory
PHENIX Collaboration

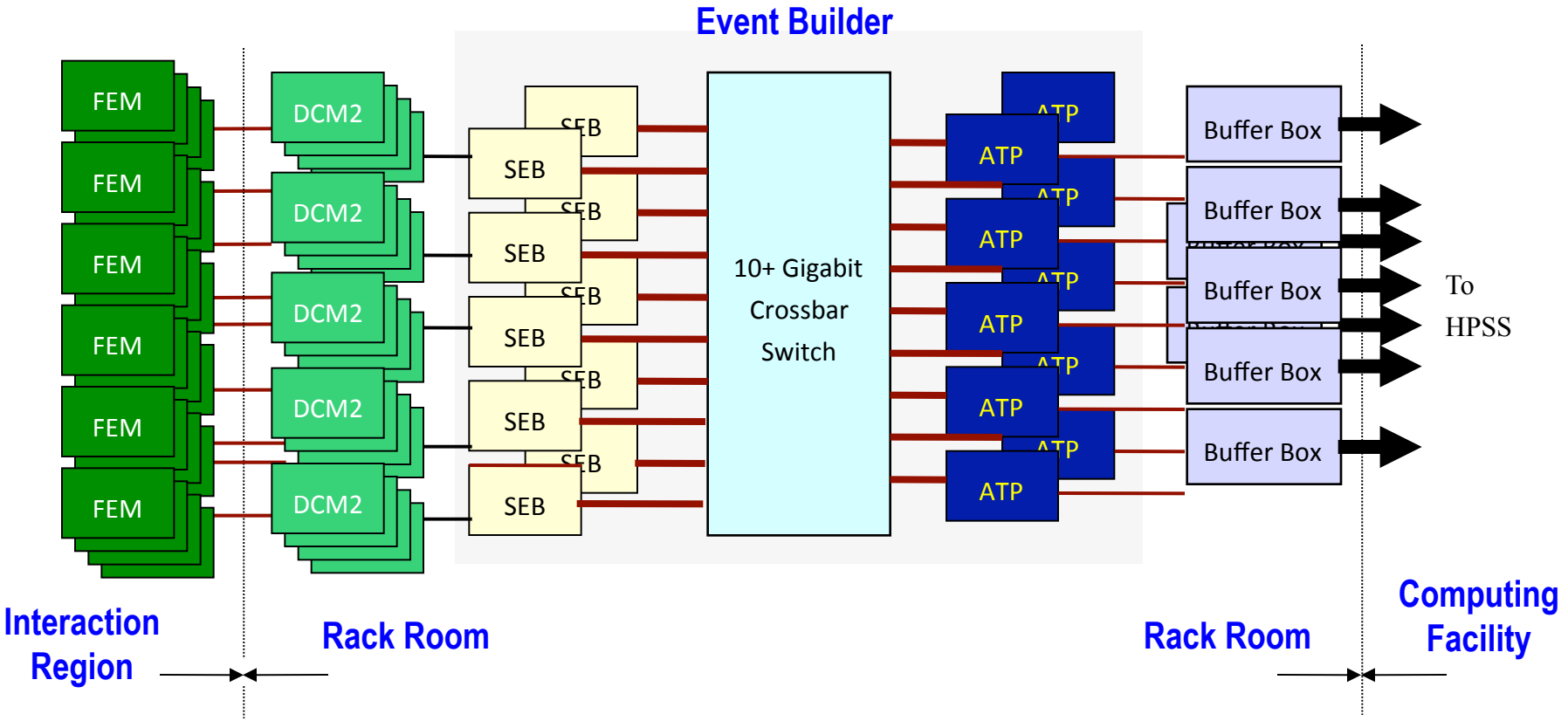
RHIC from space

Long Island, NY

1

Some Highlights…

DAQ Overview

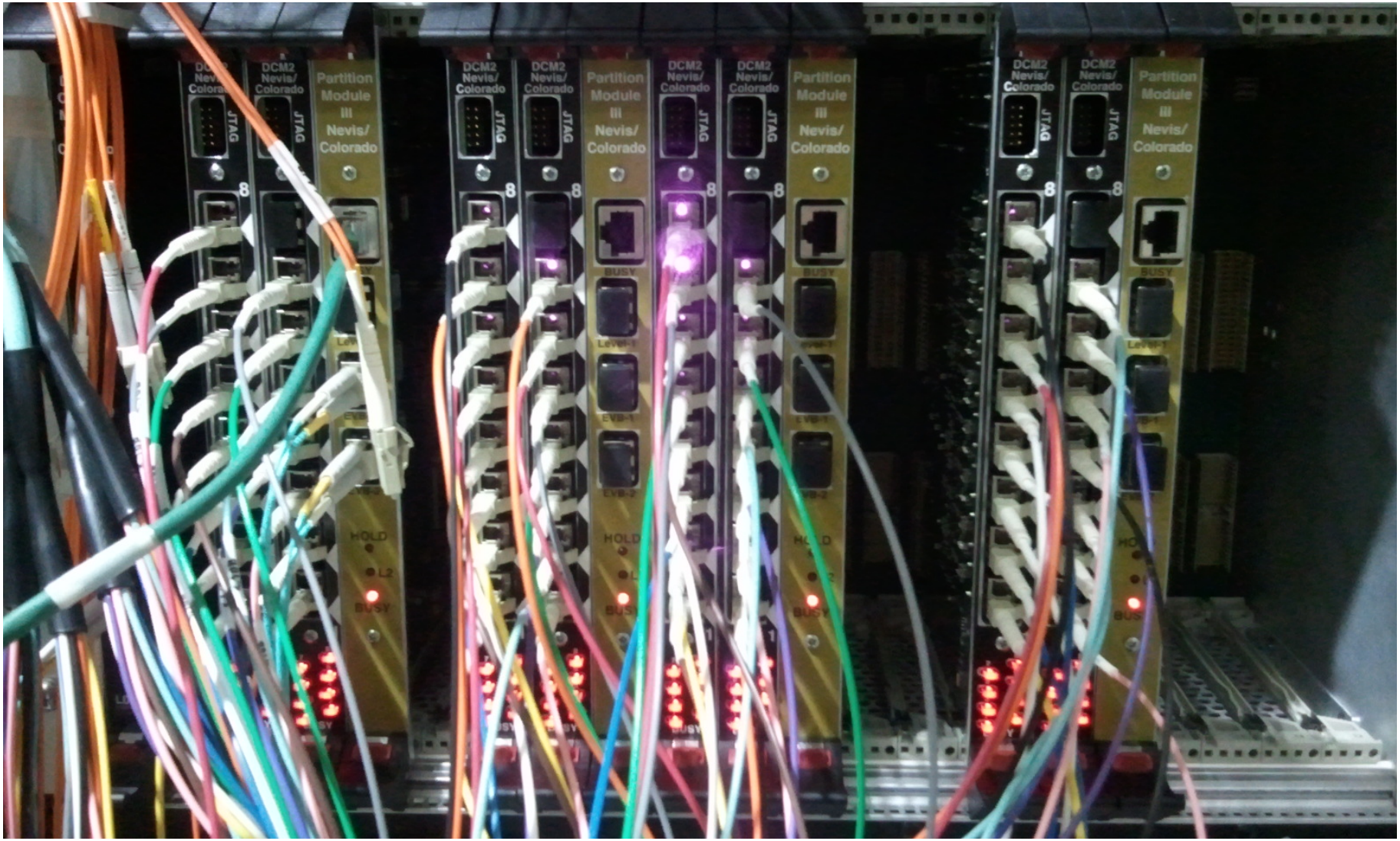


- DCM-2 receives data from digitizer, zero-suppresses and packages
- SEB collects data from a DCM group (~35)
- ATP Assembles events and compresses data (~60)
- Buffer Box data interim storage before sending to the computing center (7)

DAQ Components – DCM2

- DCM2 is a modern, mature board commissioned in PHENIX Runs 14,15
- “just as we would have built it” for sPHENIX
- FPGA code, tools, existing configurations to draw from
- Configuration tools / description language in place
- O(200) units
- Production-grade boards in hand (for test beams, R&D, etc)

DCM2's


DCM2
Nevis/
Colorado
JTAG
Partition
Module
III
Nevis/
Colorado
HOLD
L2
BUSY

Multi-Event buffering

- This is what makes the DAQ as fast as it is
- Or, in other words: that keeps the dead time in the low 90%'s
- Comes down to dealing with your data from various events in parallel

An another-world example

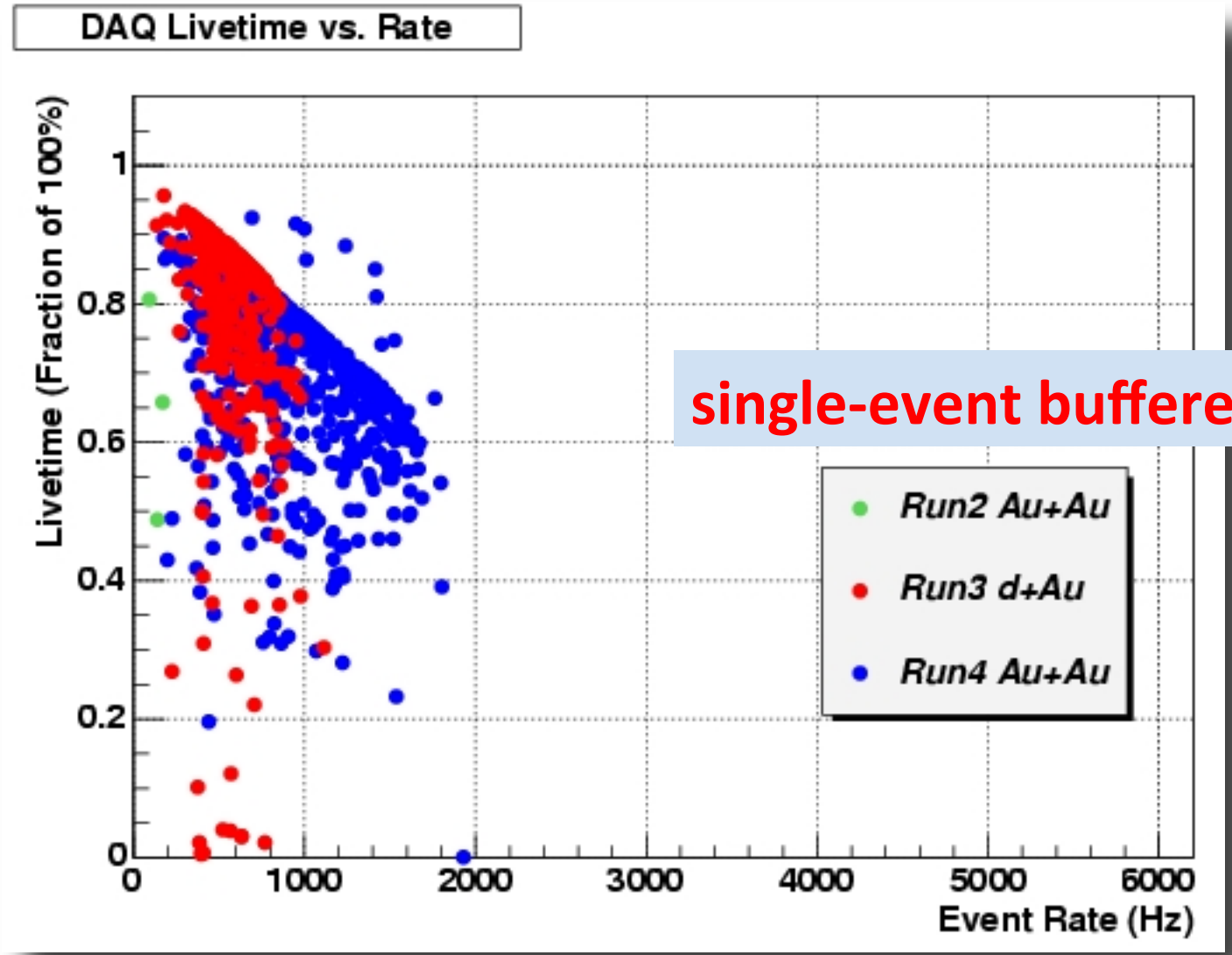
A Volkswagen assembly line

A given car takes about 28 hours from starting as a naked chassis to being an assembled vehicle

One station adds the “skin”, another the engine, another installs the defeat devices – about 340 “stations”

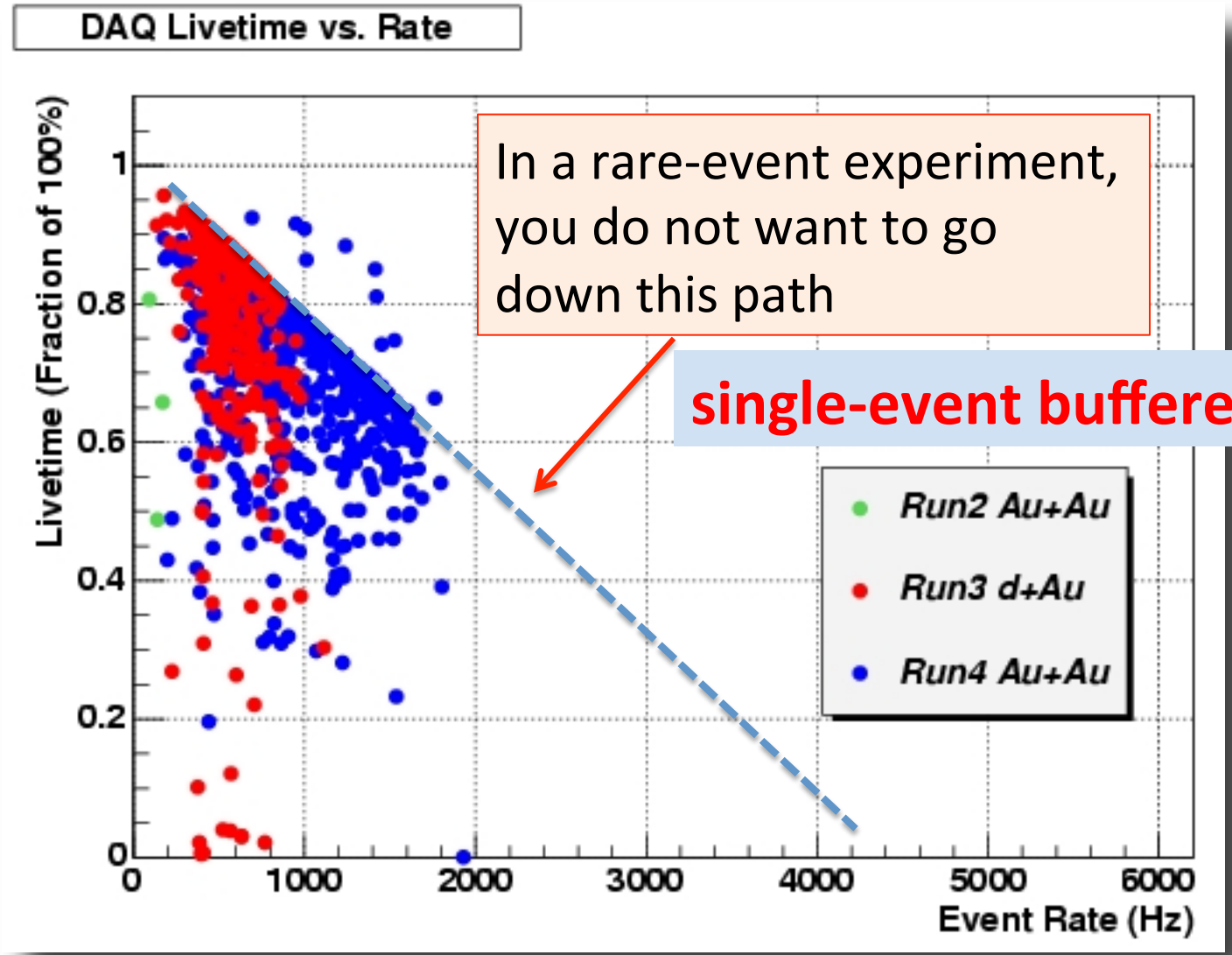
- Single-Event buffering – next car enters the assembly line once the previous car is done → **one car every 28 hours**
- Multi-Event Buffering – car moves forward as soon as the next station is free → **one car about every 5 minutes**

Multi-Event buffering effect in the DAQ



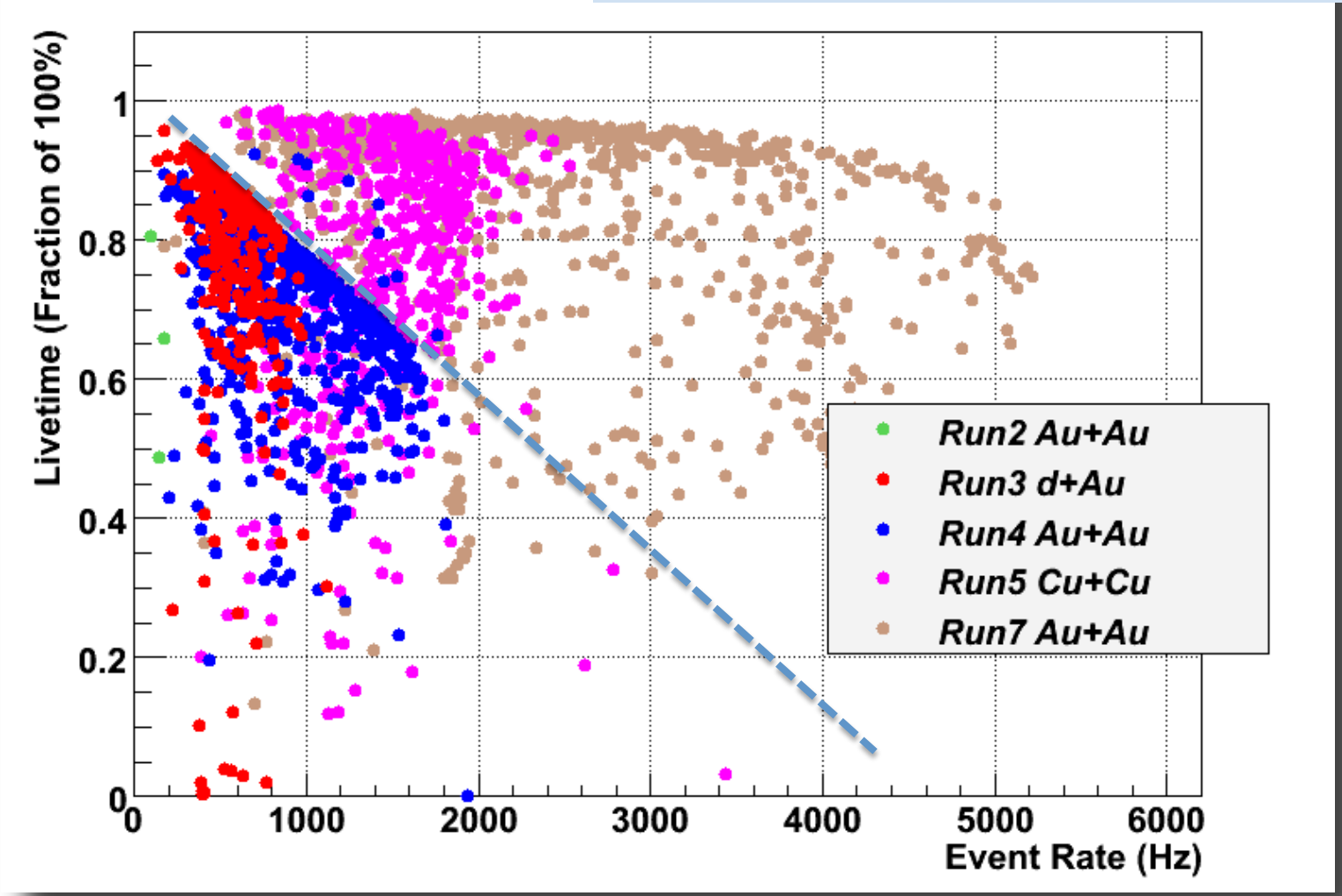
single-event buffered runs

Multi-Event buffering effect in the DAQ

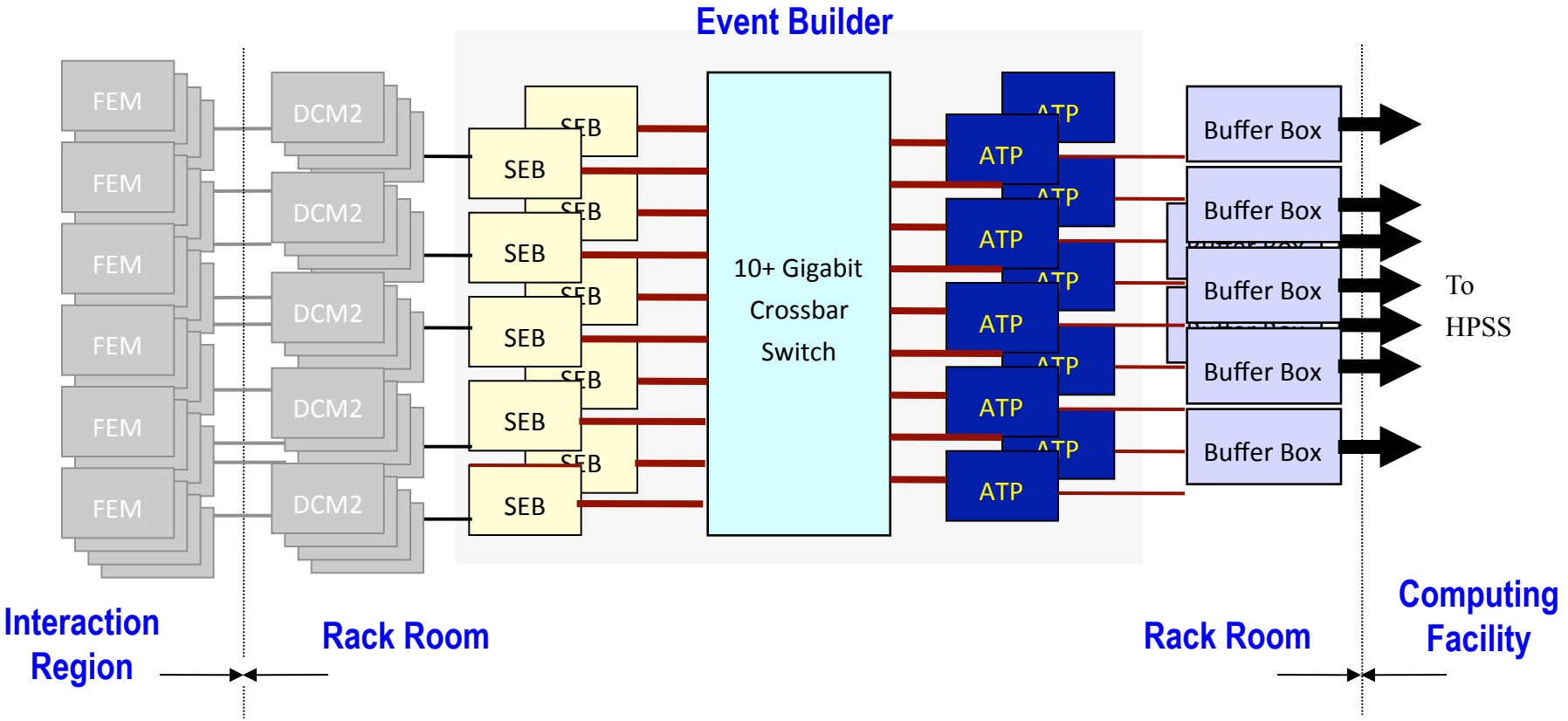


Multi-Event buffering effect in the DAQ

Plus multi-event buffered runs

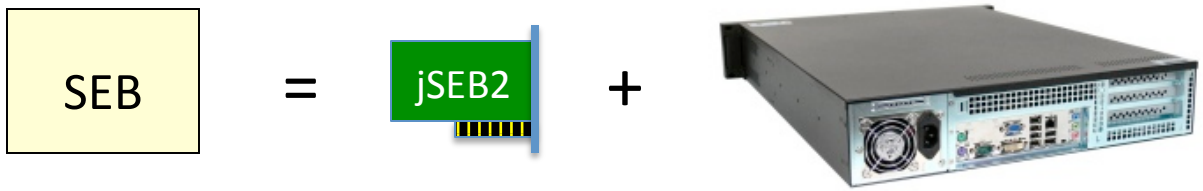


DAQ



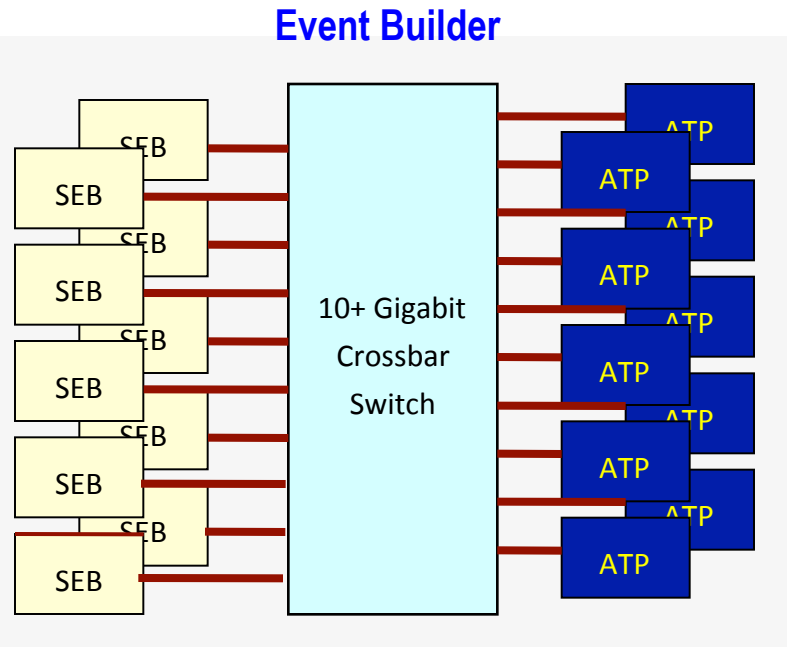
- DCM-2 receives data from digitizer, zero-suppresses and packages
- SEB collects data from a DCM group (~35)
- ATP Assembles events and compresses data (~60)
- Buffer Box data interim storage before sending to the computing center (7)

DAQ Components – jSEB2 and SEB



- A “Sub-Event Buffer” refers to a Linux PC with a “jSEB2” interface card
- Card receives data from the DCM2’s via another board
- First time data are seen in a standard PC
- jSEB2 is a 4-lane PCI-Express card, 500MByte/s-capable
- Code, tools, experience from Runs 14 & 15
- Production-grade cards in hand
- SEB is basis for test beam, R&D-type, small DAQs

Event Builder



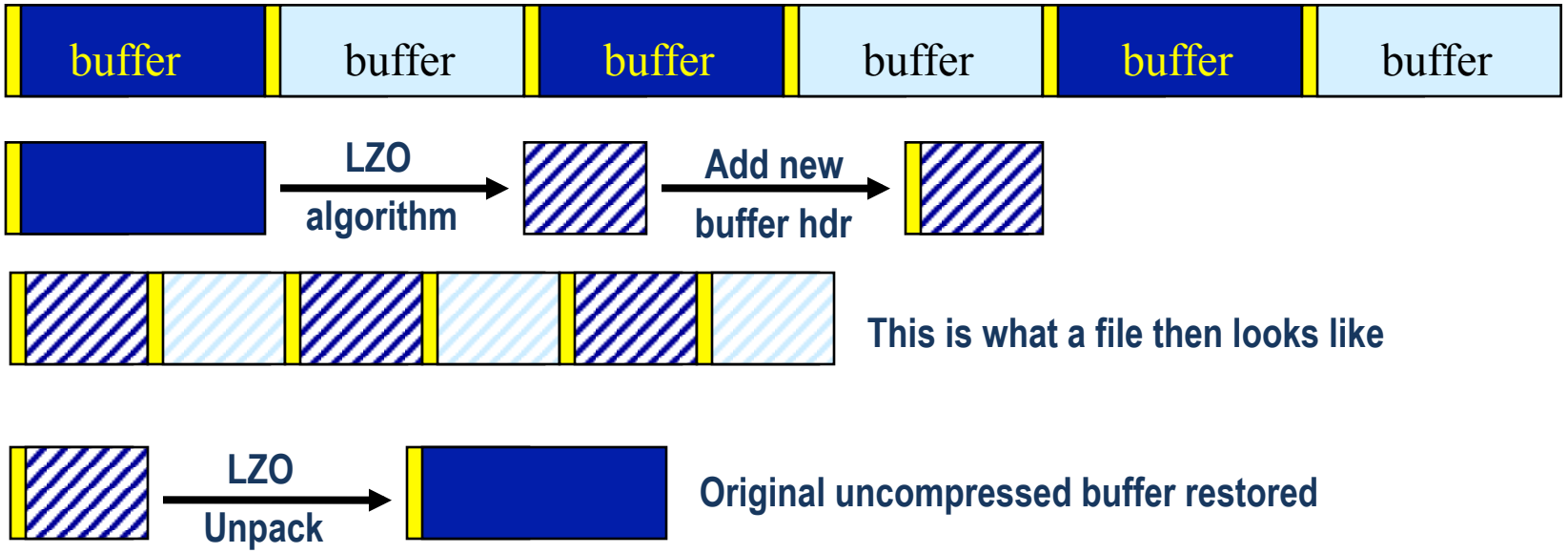
Except for ~ 20 existing machines, none of the components will be viable past run 16.
"Youngest" SEBs will support our R&D and test beam needs
File servers, buffer boxes can be re-used as needed

All PCs (SEB, ATP) are standard commercial Linux PCs
A viable, cost-efficient candidate for an EVB-grade network switch exists *today*
Larger variety of commodity-type options expected by 2020

Data compression

After all data *reduction* techniques (zero-suppression, bit-packing, etc) are applied, you typically find that your raw data are still gzip-compressible to a significant amount

Introduced a compressed raw data format that supports a late-stage compression



All this is handled completely in the I/O layer, the higher-level routines just receive a buffer as before.

Data compression load

This compression buys you a lot spare bandwidth and nice features (e.g, your analysis runs faster!)

However:

A late-stage compression (think: in your tape drive in the extreme case) doesn't really help you

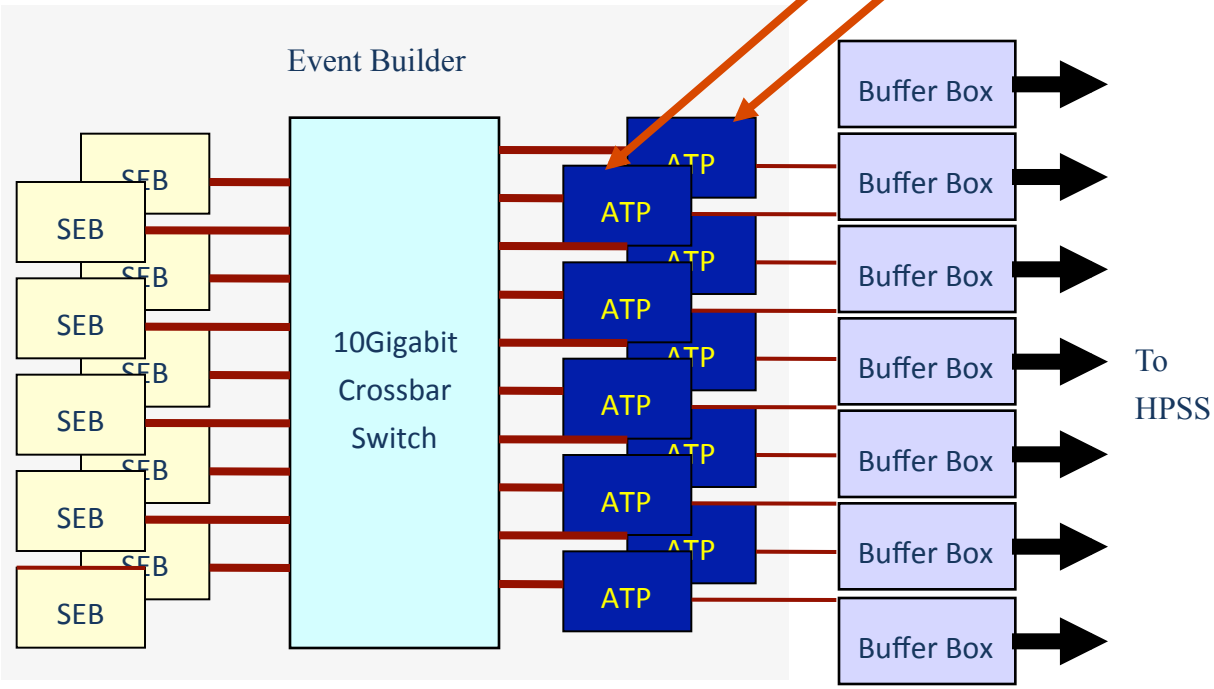
The compression has to kick in before the data hit your storage system for the first time

No single machine can keep up with compressing a >2GByte/s rate

Switch to distributed compression

The Event builder has to cope with the uncompressed data flow, e.g. 1200MB/s … 2500MB/s

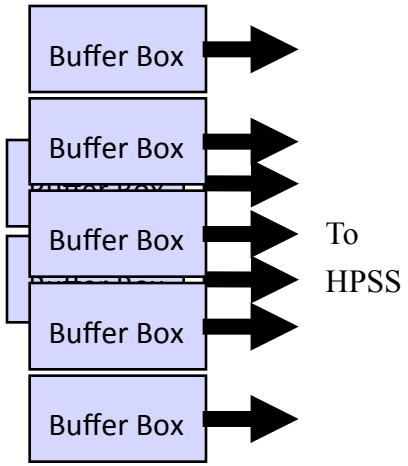
The compression is handled in the "Assembly and Trigger Processors" (ATP's) and can so be distributed over many CPU's -- that was the breakthrough



The buffer boxes and storage system see the compressed data stream, 700MB/s … 1300MB/s

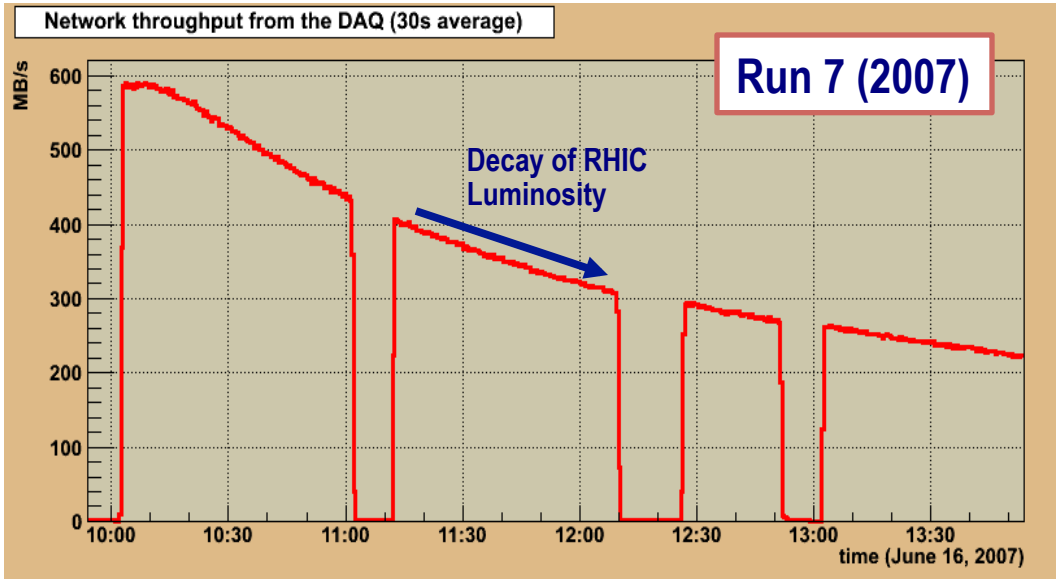
Current compression levels: 100G become ~55G
→ 45%

Network to RCF

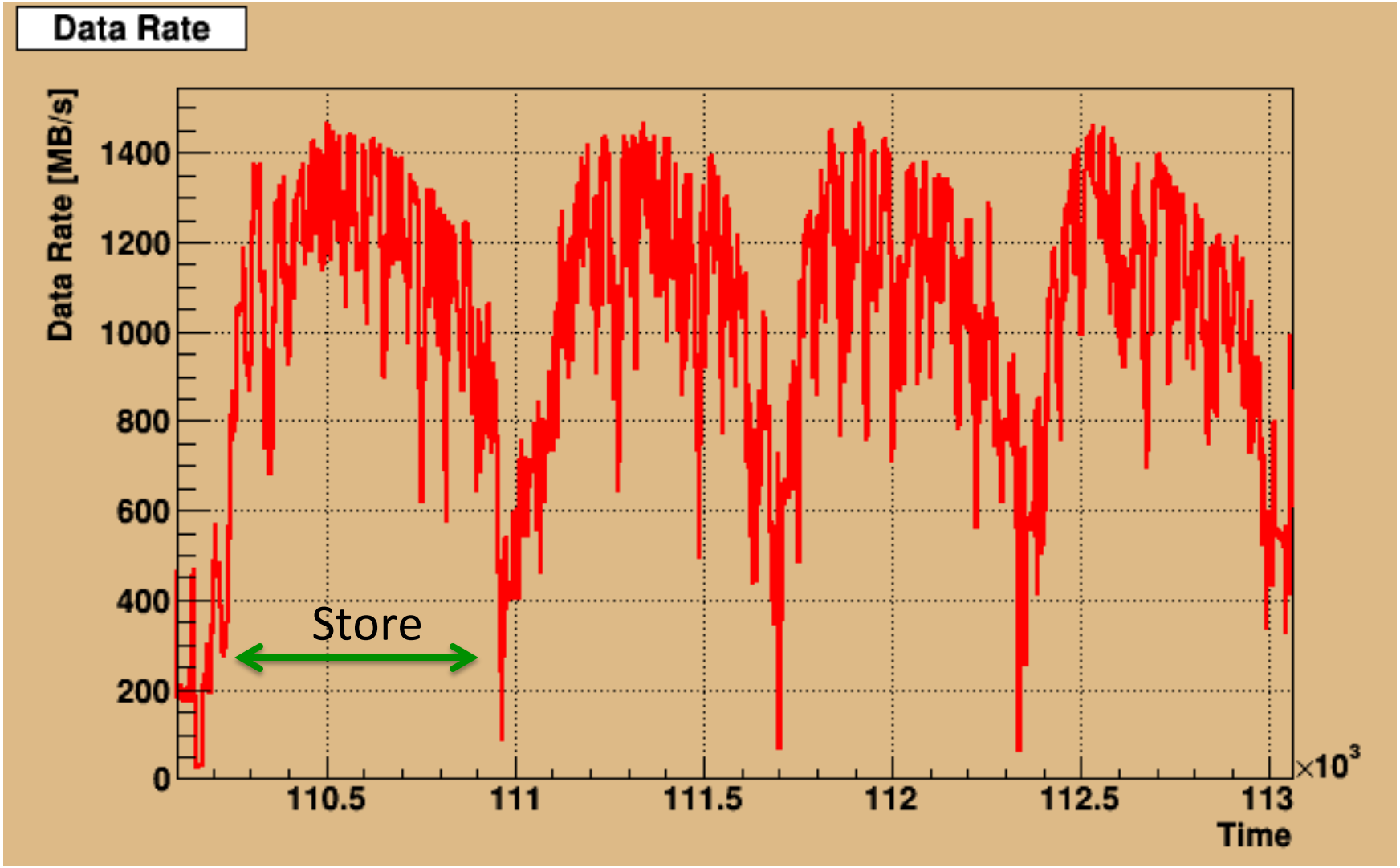


- Multiple single-mode fibers available and in production
- Capacity for 80Gbit/s configured
- Spare fibers available
- Number of buffer boxes can be tailored to match our needs

- The buffer boxes can buffer the raw data for about 80 hours
- Put in place to be able to ride out tape storage & network outages
- Allows a given file to stay in Bldg 1008 to get calibrations, 2nd-level monitoring, etc
- We only need to send the average data rates to the computing center



Data Rates (Run14)



PHENIX Raw Data Format (PRDF)

- Very mature, versatile, and “buzzword-compliant” format
- 100% payload agnostic (you can store *anything* in a PRDF – you want to add some excel spreadsheet to your raw data? An image? A drawing? No problem.)
- Support for compressed data – already discussed
- Endianess-agnostic
- Concatenation of two PRDFs yields a valid PRDF again
  (cat prdf1 prdf2 > prdfnew)
- Ready-made data inspection tools (dlist/ddump)
- Very shallow learning curve: You will see that you can analyze your LED / cosmic / test data in no time

R&D / Tests / Test beam support

- We have different DAQ systems available that all produce PRDFs; full online monitoring facilities available
- Lots of copies of a lightweight DAQ system in use in sPHENIX and the EIC orbit (BNL, UIUC, SBU, Yale, GSU, even ATLAS ZDC test beam/calibration)
- You can take all your R&D-related data in PRDF format from the get-go
- No hard switchover once you move to the “big” sPHENIX DAQ
- Lots of commodity hardware you might have supported out of the box – DRS4 boards, SRS, Struck FADCs, etc etc

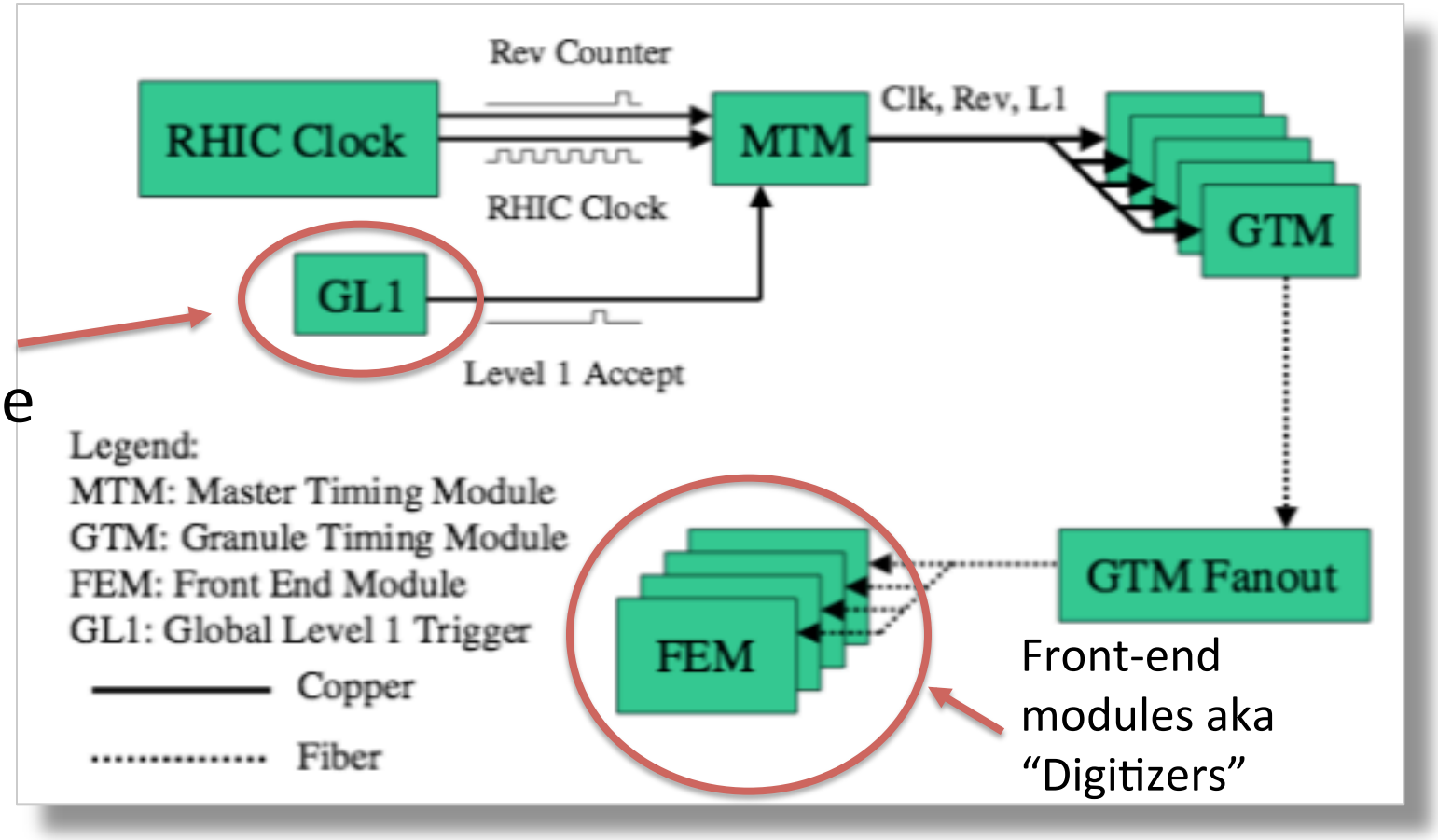
Issues

- Potential TPC tracking system ("continuous readout") needs to be integrated into the DAQ
- Continuous readout doesn't follow event paradigm
- Work on Event builder needed to accommodate TPC data

(I have been talking to groups who do similar things – concept not that complicated, but the details are)

Trigger Electronics & Timing system

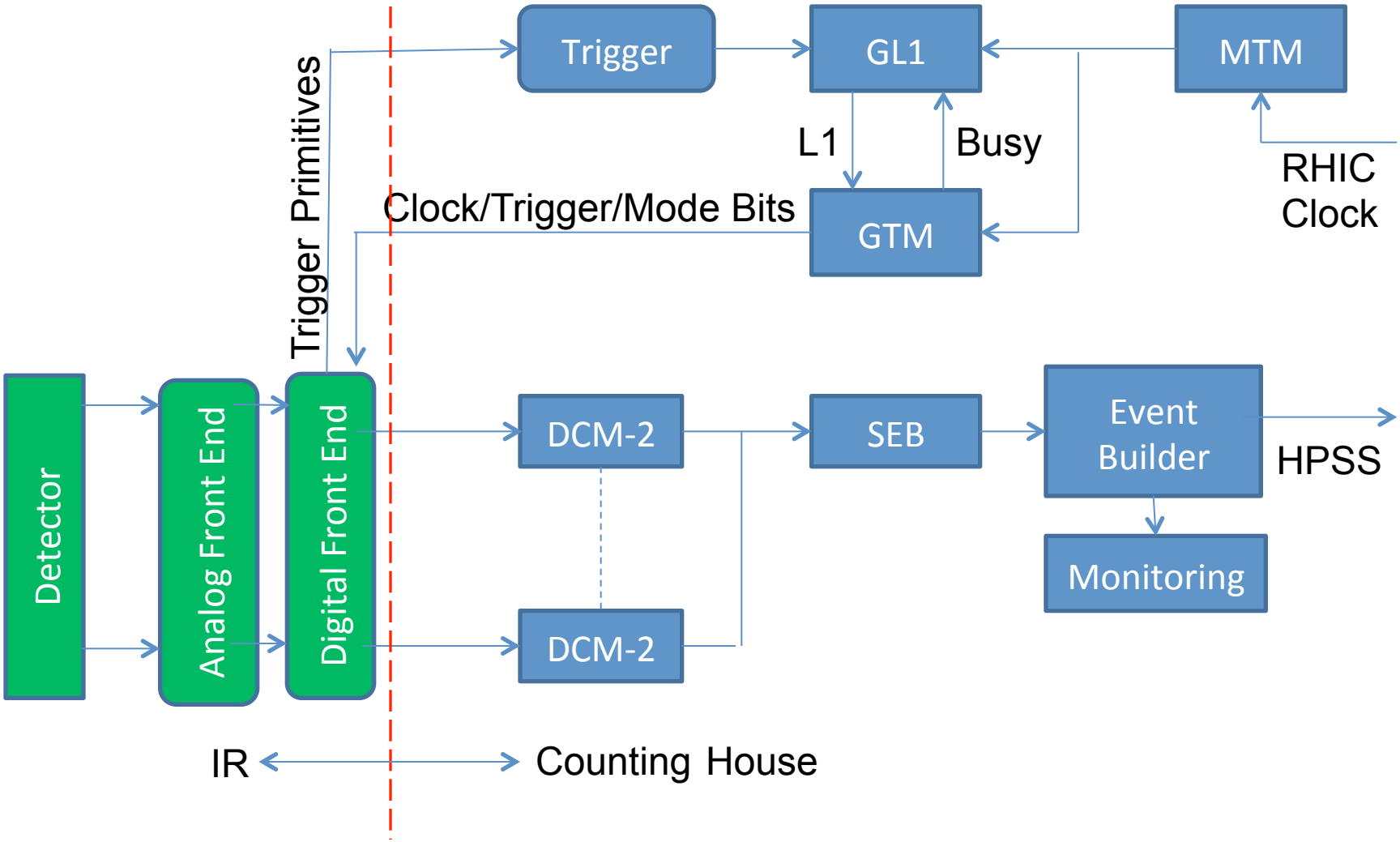
Trigger interface



Local-Level 1 (LL1) Boards

- LL1's generate a trigger signal from trigger primitives generated in the front-end
- new LL1 boards for calorimeter triggers
- Global L1 – upgrade scalers? Current 32bit scalers roll over
- Potential new triggers for p+p, p+A need LL1 boards.

Trigger & Timing system



Trigger

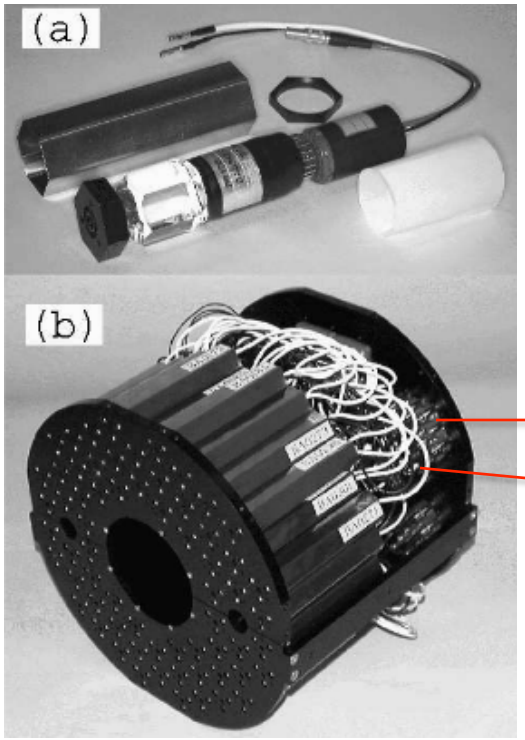
- Mostly “minimum bias” for A+A
- Min. Bias + dedicated triggers for p+p, p+A
- Log all that passes Lvl1 trigger (no lvl2-2 trigger)
- Slightly higher data volume buys physics one cannot readily trigger on (e.g. Heavy Flavor, ultraperipheral)
- Good experience with this approach in PHENIX

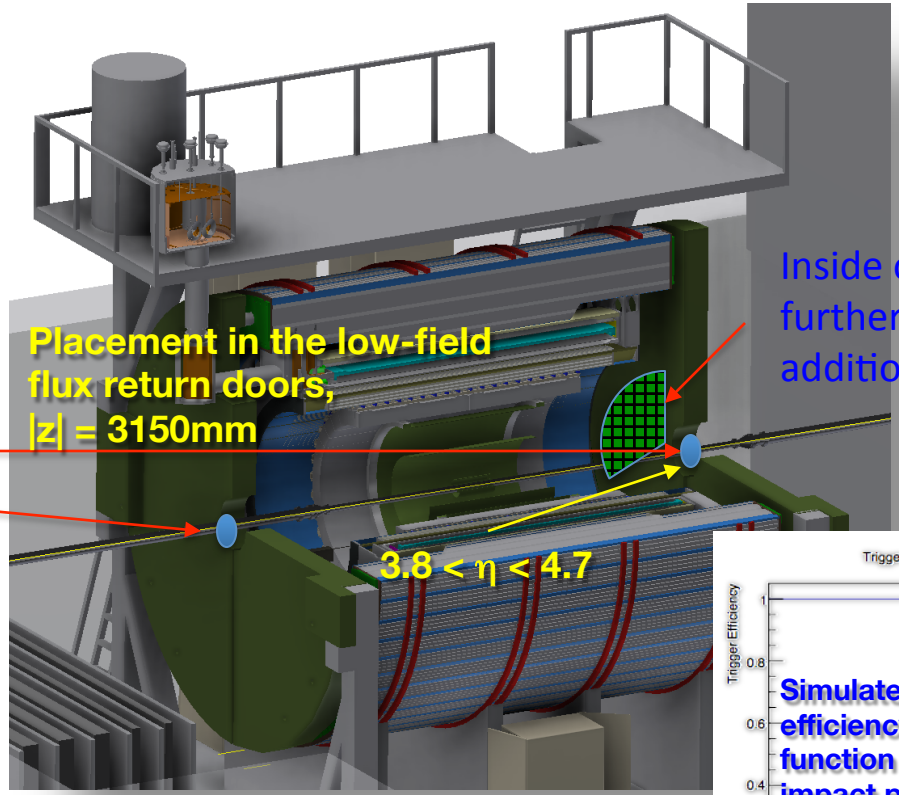
Trigger Detectors

- detector to “recognize” a collision
- Timing signal o(50ps) for vertex determination
- Good timing (Cherenkov, photomultipliers) and magnetic fields don’t mix well
- One timing option would be to recycle the existing PHENIX Beam-Beam Counter, placed outside the field
- Add a “non-timing” detector for more coverage in the field

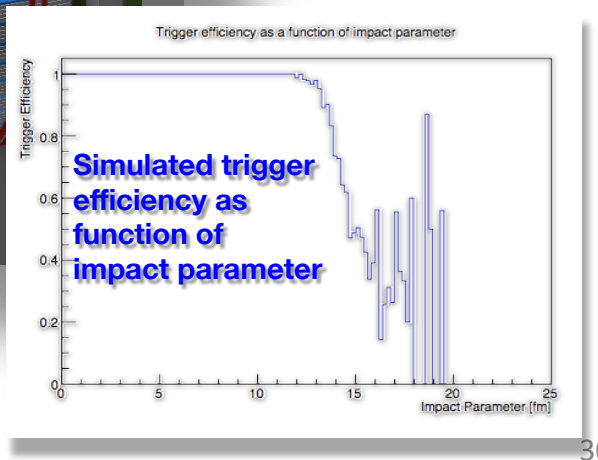
PHENIX BBC Option

- 2x (north and south) 64 3cm quartz Cherenkov radiators
- 2x 64 one-inch diameter mesh-dynode photomultiplier tubes (Hamamatsu R6178





Inside of the doors (or further in) can be used for additional coverage



BBC Alternatives

- The current BBC in hand would be sufficient to provide the timing
- If we decide to rebuild the BBC, we might further optimize the parameters (larger crystals/more coverage)
- Additional detector would be unchanged
- Would act as a good reaction plane detector offline
- Simulations of various scenarios ongoing (slowly)

BNL cannot take this on

Great opportunity for another group to jump in

Issues and concerns - Trigger

- Trigger detector (BBC) needs more simulations and design work
- Unlikely that the PHENIX BBC will be available
- Design for the “inner-door” part needs work
- Also open for other, not drawn-on-a-napkin concepts
- Groups to take on these parts need to be identified/come forward
- Electronic boards to be built
- 64bit upgrade of all 32bit scalers